Caspian Journal of Applied Mathematics, Ecology and Economics
V. 2, No 1, 2014, July
ISSN 1560-4055

A Type of Shannon-McMillan Approximation Theorems for Second-Order Nonhomogeneous Markov Chains Indexed by a Double Rooted Tree

Wang Kangkang *, Zong Decai

Abstract. In this paper, a class of small deviation theorems for the relative entropy densities of arbitrary random field on a double rooted tree are discussed by comparing between the arbitrary measure μ and the second-order nonhomogeneous Markov measure μ_Q on the double rooted tree. As corollaries, some Shannon-McMillan theorems for the arbitrary random field, second-order Markov chain field and a limit property for the random conditional entropy of second-order homogeneous Markov chain on the double rooted tree are obtained. The existing result is extended.

Key Words and Phrases: Shannon-McMillan theorem, a double rooted tree, arbitrary random field, second-order nonhomogeneous Markov chain, relative entropy density.
2000 Mathematics Subject Classifications: 60F15

---

1. Introduction

A tree is a graph $S = \{T, E\}$ which is connected and contains no circuits. Given any two vertices $\sigma, t(\ \sigma \neq t \in T)$, let $\overline{\sigma t}$ be the unique path connecting σ and t. Define the graph distance $d(\sigma, t)$ to be the number of edges contained in the path $\overline{\sigma t}$.

Let T_o be an arbitrary infinite tree that is partially finite (i.e. it has infinite vertices, and each vertex connects with finite vertices) and has a root o. Meanwhile, we consider another kind of double root tree T, that is, it is formed with the root o of T_o connecting with an arbitrary point denoted by the root -1. For a better explanation of the double root tree T, we take Cayley tree $T_{C,N}$ for example. It's a special case of the tree T_o, the root o of Cayley tree has N neighbors and all the other vertices of it have $N+1$ neighbors each. The double root tree $T'_{C,N}$ (see Fig.1) is formed with root o of tree $T_{C,N}$ connecting with another root -1.

Let σ, t be vertices of the double root tree T. Write $t \leq \sigma$ $(\sigma, t \neq o, -1)$ if t is on the unique path connecting o to σ, and $|\sigma|$ for the number of edges on this path. For any

*Corresponding author.

two vertices σ, t ($\sigma, t \neq o, -1$) of the tree T, denote by $\sigma \wedge t$ the vertex farthest from o satisfying $\sigma \wedge t \leq \sigma$ and $\sigma \wedge t \leq t$.

The set of all vertices with distance n from root o is called the n-th generation of T, which is denoted by L_n. We say that L_n is the set of all vertices on level n and especially root -1 is on the -1st level on tree T. We denote by $T^{(n)}$ the subtree of the tree T containing the vertices from level -1 (the root -1) to level n and denote by $T_o^{(n)}$ the subtree of the tree T_o containing the vertices from level 0 (the root o) to level n. Let $t(\neq o, -1)$ be a vertex of the tree T. We denote the first predecessor of t by 1_t, the second predecessor of t by 2_t, and denote by n_t the n-th predecessor of t. Let $X^A = \{X_t, t \in A\}$, and let x^A be a realization of X^A and denote by $|A|$ the number of vertices of A .

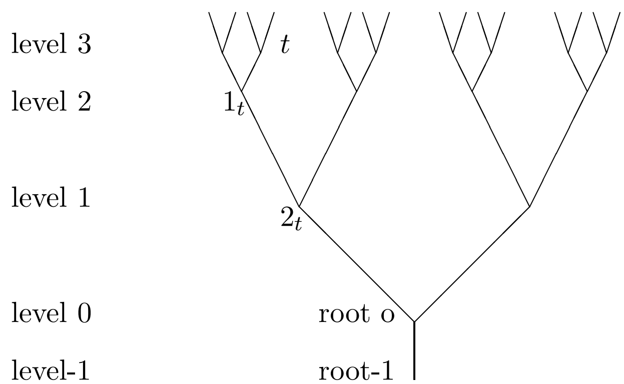

Fig.1 Double root tree $T'_{C,2}$

Definition 1 *Let $S = \{s_1, s_2, \cdots, s_M\}$ and $Q(z|y,x)$ be a nonnegative function on S^3. Let*

$$Q = ((Q(z|y,x)), \quad Q(z|y,x) \geq 0, \quad x, y, z \in S.$$

If

$$\sum_{z \in S} Q(z|y,x) = 1,$$

then Q is called a second-order transition matrix.

Definition 2 *Let T be a double root tree and $S = \{s_1, s_2, \cdots, s_M\}$ be a finite state space, and $\{X_t, t \in T\}$ be a collection of S-valued random variables defined on the probability space $(\Omega, \mathcal{F}, Q)$. Let*

$$q = (q(x,y)), \quad x, y \in S \tag{1}$$

be a distribution on S^2, and

$$Q_t = (Q_t(z|y,x)), \quad x, y, z \in S, \quad t \in T \backslash \{o\}\{-1\} \tag{2}$$

be a collection of second-order transition matrices. For any vertex t ($t \neq o, -1$), if

$$Q(X_t = z|X_{1_t} = y, X_{2_t} = x, \text{ and } X_\sigma \text{ for } \sigma \wedge t \leq 2_t)$$

$$= Q(X_t = z|X_{1_t} = y, X_{2_t} = x) = Q_t(z|y, x) \quad \forall x, y, z \in S \tag{3}$$

and

$$Q(X_{-1} = x, X_o = y) = q(x, y), \quad x, y \in S, \tag{4}$$

then $\{X_t, t \in T\}$ is called a S-valued second-order nonhomogeneous Markov chain indexed by a tree T with the initial distribution (1) and second-order transition matrices (2), or called a T-indexed second-order nonhomogeneous Markov chain.

Definition 3. *Let $(Q_t = Q_t(z|x, y), t \in T^{(n)} \backslash \{o, -1\})$ and $q = (q(x, y))$ be defined as before, μ_Q be a second-order nonhomogeneous Markov measure on $(\Omega, \mathcal{F})$. If*

$$\mu_Q(x_0, x_{-1}) = q(x_0, x_{-1}) \tag{5}$$

$$\mu_Q(x^{T^{(n)}}) = q(x_0, x_{-1}) \prod_{t \in T^{(n)} \backslash \{o, -1\}} Q_t(x_t|x_{1_t}, x_{2_t}), \quad n \geq 1, \tag{6}$$

then μ_Q will be called a second-order Markov chains field on an infinite tree T determined by the stochastic matrices Q_t and the initial distribution q.

Let μ be an arbitrary probability measure defined on $(\Omega, \mathcal{F})$, *log* is the natural logarithmic. Denote

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \log \mu(X^{T^{(n)}}). \tag{7}$$

$f_n(\omega)$ is called the entropy density on subgraph $T^{(n)}$ with respect to the measure μ. If $\mu = \mu_Q$, then by (6), (7) we get

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}[\log q(X_0, X_{-1}) + \sum_{t \in T^{(n)} \backslash \{o, -1\}} Q_t(X_t|X_{1_t}, X_{2_t})]. \tag{8}$$

The convergence of $f_n(\omega)$ in a sense (L_1 convergence, convergence in probability, or almost sure convergence) is called Shannon-McMillan theorem or the asymptotic equipartition property (AEP) in information theory. There have been some works on limit theorems for tree-indexed stochastic processes. Benjamini and Peres [1] have given the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [2] have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Ye and Berger (see [4],[5]), by using Pemantle's result [3] and a combinatorial approach, have studied the Shannon-McMillan theorem with convergence in probability for a PPS-invariant and ergodic random field on a homogeneous tree. Yang and Liu [8] have studied a strong law of large numbers for the frequency of occurrence of states for Markov chains field on a homogeneous tree (a particular case of tree-indexed Markov chains field and PPS-invariant random fields). Yang (see [6]) has studied the strong law of large numbers for frequency of occurrence of state and Shannon-McMillan theorem for homogeneous Markov chains indexed by a homogeneous tree. Recently, Yang (see [13]) has studied the strong law of large numbers and Shannon-McMillan theorem for nonhomogeneous Markov chains indexed by a homogeneous tree. Huang and Yang (see [11]) have also studied the strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. Peng and Yang have studied a class of small deviation theorems for functionals for arbitrary random field on a homogeneous trees (see[9]). Wang has also studied some Shannon-McMillan approximation theorems for arbitrary random field on the generalized Bethe tree (see[10]).

In this paper, we study a class of Shannon-McMillan random approximation theorems for arbitrary random fields on the double rooted tree by comparison the arbitrary measure with the second-order nonhomogeneous Markov measure and constructing a super-martingale on the double rooted tree. As corollaries, a class of Shannon-McMillan theorems for arbitrary random field and second-order Markov chains field on the double rooted tree are obtained. A limit property for the expectation of the random conditional entropy of second-order homogeneous Markov chain indexed by the double rooted tree is studied. Yang and Ye's result (see[13]) is extended.

2. Main result and its proof

Lemma 1 (see [8]). *Let μ_1 and μ_2 be two probability measures defined on $(\Omega, \mathcal{F})$, $D \in \mathcal{F}$, $\{\tau_n, n \geq 0\}$ be a sequence of positive-valued random variables such that*

$$\liminf_n \frac{\tau_n}{|T^{(n)}|} > 0. \quad \mu_1 - \quad a.s. \quad D. \tag{10}$$

Then

$$\limsup_{n\to\infty} \frac{1}{\tau_n} \log \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0. \quad \mu_1 - \quad a.s. \quad D. \tag{11}$$

In particular, let $\tau_n = |T^{(n)}|$, then

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0. \qquad \mu_1 - \quad a.s. \tag{12}$$

Proof . See reference [8].

Let

$$\varphi(\mu|\mu_Q) = \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})}, \tag{13}$$

$\varphi(\mu|\mu_Q)$ is called the sample relative entropy rate of $X^{T^{(n)}}$ with respect to μ and μ_Q. $\varphi(\mu|\mu_Q)$ is also called asymptotic logarithmic likelihood ratio. By (12) and (13)

$$\varphi(\mu|\mu_Q) \geq \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})} \geq 0. \quad \mu - a.s. \tag{14}$$

Hence $\varphi(\mu|\mu_Q)$ can be looked on as a type of a measure of the deviation between the arbitrary random field and the second-order nonhomogeneous Markov chain fields on the double rooted tree.

Although $\varphi(\mu|\mu_Q)$ is not a proper metric between two probability measures, we nevertheless think of it as a measure of "dissimilarity" between their joint distribution μ and second-order Markov distribution μ_Q. Obviously, $\varphi(\mu|\mu_Q) = 0$ if and only if $\mu = \mu_Q$. It has been shown in (14) that $\varphi(\mu|\mu_Q) \geq 0$, a.s. in any case. Hence, $\varphi(\mu|\mu_Q)$ can be used as a random measure of the deviation between the true joint distribution $\mu(x^{T^{(n)}})$ and the second-order Markov distribution $\mu_Q(x^{T^{(n)}})$. Roughly speaking, this deviation may be regarded as the one between coordinate stochastic process $x^{T^{(n)}}$ and the Markov case. The smaller $\varphi(\mu|\mu_Q)$ is, the smaller the deviation is.

Theorem 1. *Let $X = \{X_t, t \in T\}$ be an arbitrary random field on a double rooted tree. $f_n(\omega)$ and $\varphi(\mu|\mu_Q)$ are respectively defined as (7) and (13). Denote by $H_t^Q(X_t|X_{1_t}, X_{2_t})$ the random conditional entropy of X_t relative to X_{1_t}, X_{2_t} on the measure μ_Q, that is*

$$H_t^Q(X_t|X_{1_t}, X_{2_t}) = -\sum_{x_t\in S} Q_t(x_t|X_{1_t}, X_{2_t}) \log Q_t(x_t|X_{1_t}, X_{2_t}), \qquad t \in T^{(n)}\backslash\{o, -1\}. \tag{15}$$

Let

$$D(c) = \{\omega : \varphi(\mu|\mu_Q) \leq c\} \tag{16}$$

$$\alpha(c) = \min\left\{\frac{2xe^{-2}}{(1-x)^2} + \frac{c}{x}, \quad 0 < x < 1\right\}, \quad c > 0; \quad \alpha(0) = 0. \tag{17}$$

$$\beta(c) = \max\left\{\frac{2xe^{-2}}{(1+x)^2} + \frac{c}{x}, \quad -1 < x < 0\right\}, \quad c > 0; \quad \beta(0) = 0. \tag{18}$$

Then

$$\limsup_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\backslash\{o,-1\}} H_t^Q(X_t|X_{1_t}, X_{2_t})] \leq \alpha(c)M, \quad \mu - a.s. \ \ \omega \in D(c) \tag{19}$$

$$\liminf_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\backslash\{o,-1\}} H_t^Q(X_t|X_{1_t}, X_{2_t})] \geq \beta(c)M - c. \quad \mu - a.s. \ \ \omega \in D(c) \tag{20}$$

Proof. Consider the probability space $(\Omega, \mathcal{F}, \mu)$, let $\lambda > 0$ be a constant, $\delta_j(\cdot)$ be Kronecker function. Denote $g_t(j) = -\log Q_t(j|X_{1_t}, X_{2_t})$, we construct the following product distribution:

$$\mu_Q(x^{T^{(n)}};\lambda) = q(x_0, x_{-1}) \prod_{t\in T^{(n)}\setminus\{o,-1\}} \exp\{\lambda g_t(j)\delta_j(x_t)\}[\frac{Q_t(x_t|x_{1_t}, x_{2_t})}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}]. \tag{22}$$

By (22) we can write

$$\begin{aligned}
&\sum_{x^{L_n}\in S^{L_n}} \mu_Q(x^{T^{(n)}};\lambda)\\
&= \sum_{x^{L_n}\in S^{L_n}} q(x_0, x_{-1}) \prod_{t\in T^{(n)}\setminus\{o,-1\}} \exp\{\lambda g_t(j)\delta_j(x_t)\}[\frac{Q_t(x_t|x_{1_t}, x_{2_t})}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}]\\
&= \mu_Q(x^{T^{(n-1)}};\lambda) \sum_{x^{L_n}\in S^{L_n}} \prod_{t\in L_n} \exp\{\lambda g_t(j)\delta_j(x_t)\}[\frac{Q_t(x_t|x_{1_t}, x_{2_t})}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}]\\
&= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{t\in L_n} \sum_{x_t\in S} \exp\{\lambda g_t(j)\delta_j(x_t)\}[\frac{Q_t(x_t|x_{1_t}, x_{2_t})}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}]\\
&= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{t\in L_{n-1}} \frac{1}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}[\sum_{x_t=j} + \sum_{x_t\neq j}]\\
&= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{t\in L_{n-1}} \frac{e^{\lambda g_t(j)}Q_t(j|x_{1_t}, x_{2_t}) + 1 - Q_t(j|x_{1_t}, x_{2_t})}{1+(e^{\lambda g_t(j)}-1)Q_t(j|x_{1_t}, x_{2_t})}\\
&= \mu_Q(x^{T^{(n-1)}};\lambda)
\end{aligned} \tag{23}$$

Therefore $\mu_Q(x^{T^{(n)}};\lambda)$, $n = 1, 2, \cdots$ are a class of consistent distributions on $S^{T^{(n)}}$. Let

$$U_n(\lambda, \omega) = \frac{\mu_Q(X^{T^{(n)}};\lambda)}{\mu(X^{T^{(n)}})} \tag{24}$$

By (22) and (24) we attain

$$\begin{aligned}
U_n(\lambda, \omega) = \exp\{&\sum_{t\in T^{(n)}\setminus\{o,-1\}} \lambda g_t(j)\delta_j(X_t)\} \prod_{t\in T^{(n)}\setminus\{o,-1\}} [\frac{1}{1+(e^{\lambda g_t(j)}-1)Q_t(j|X_{1_t}, X_{2_t})}]\\
&\cdot q(X_0) \prod_{t\in T^{(n)}\setminus\{o,-1\}} Q_t(X_t|X_{1_t}, X_{2_t}) \Big/ \mu(X^{T^{(n)}}).
\end{aligned} \tag{25}$$

It is easy to see that $U_n(\lambda, \omega)$ is a nonnegative sup-martingale from Doob's martingale convergence theorem (see [12]) since μ and μ_Q are two probability measures. Moreover,

$$\lim_{n\to\infty} U_n(\lambda, \omega) = U_\infty(\lambda, \omega) < \infty. \quad \mu - a.s. \tag{26}$$

By (12) and (24) we have

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log U_n(\lambda,\omega) \le 0. \qquad \mu - a.s. \tag{27}$$

According to (6), (25), we can rewrite (27) as

$$\limsup_{n\to\infty}\{\frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \lambda g_t(j)\delta_j(X_t)$$
$$-\frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \log[1+(e^{\lambda g_t(j)}-1)Q_t(j|X_{1_t},X_{2_t})] + \frac{1}{|T^{(n)}|}\log\frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}\}$$
$$\le 0 \qquad \mu - a.s. \tag{28}$$

Letting $\lambda = 0$ in (28), we have

$$\varphi(\mu|\mu_Q) \ge \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \log\frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})} \ge 0. \qquad \mu - a.s. \qquad \omega \in D(c). \tag{29}$$

By use of (16) and (28) we obtain

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \{\lambda g_t(j)\delta_j(X_t) - \log[1+(e^{\lambda g_t(j)}-1)Q_t(j|X_{1_t},X_{2_t})\}$$
$$\le \varphi(\mu|\mu_Q) \le c. \qquad \mu - a.s. \qquad \omega \in D(c). \tag{30}$$

By virtue of (30), the properties of super limit and the inequalities $1-1/x \le \ln x \le x-1, (x>0)$, $e^x - 1 - x \le (1/2)x^2 e^{|x|}$, we can write

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \lambda\{g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t},X_{2_t})\}$$

$$\le \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \{\log[1+(e^{\lambda g_t(j)}-1)Q_t(j|X_{1_t},X_{2_t})] - \lambda g_t(j)Q_t(j|X_{1_t},X_{2_t})\}+c$$

$$\le \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} Q_t(j|X_{1_t},X_{2_t})[e^{\lambda g_t(j)} - 1 - \lambda g_t(j)] + c$$

$$\le (\lambda^2/2)\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} Q_t(j|X_{1_t},X_{2_t}) g_t^2(j) e^{|\lambda g_t(j)|} + c$$

$$= (\lambda^2/2)\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} Q_t(j|X_{1_t},X_{2_t}) \log^2 Q_t(j|X_{1_t},X_{2_t}) e^{-|\lambda|\log Q_t(j|X_{1_t},X_{2_t})}+c$$

$$= (\lambda^2/2)\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log^2 Q_t(j|X_{1_t},X_{2_t})\cdot Q_t(j|X_{1_t},X_{2_t})^{1-|\lambda|}+c.$$

$$\mu-a.s.\qquad \omega\in D(c) \tag{31}$$

In the case $0<\lambda<1$, dividing two sides of (31) by λ, we have

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t)-g_t(j)Q_t(j|X_{1_t},X_{2_t})]\\ \le\ &\frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log^2 Q_t(j|X_{1_t},X_{2_t})\cdot Q_t(j|X_{1_t},X_{2_t})^{1-\lambda}+\frac{c}{\lambda}\end{aligned}$$

$$\mu-a.s.\qquad \omega\in D(c) \tag{32}$$

Consider the function

$$\phi(x)=(\log x)^2x^{1-\lambda},\quad 0<x\le 1,\quad 0<\lambda<1.\ \ (set\ \ \phi(0)=0) \tag{33}$$

It can be concluded that on the internal $[0,1]$,

$$\max\{\phi(x),0\le x\le 1\}=\phi(e^{2/(\lambda-1)})=(\frac{2}{\lambda-1})^2e^{-2}. \tag{34}$$

By (32) and (34) we have that when $0<\lambda<1$,

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t)-g_t(j)Q_t(j|X_{1_t},X_{2_t})]\\ \le\ &\frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}(\frac{2}{\lambda-1})^2e^{-2}+\frac{c}{\lambda}\\ =\ &\frac{2\lambda e^{-2}}{(1-\lambda)^2}\limsup_{n\to\infty}\frac{|T^{(n)}|-2}{|T^{(n)}|}+\frac{c}{\lambda}\le\frac{2\lambda e^{-2}}{(1-\lambda)^2}+\frac{c}{\lambda}.\qquad \mu-a.s.\qquad \omega\in D(c)\end{aligned} \tag{35}$$

When $c>0$, $h(\lambda)=(2\lambda e^{-2})/(1-\lambda)^2$+c/$\lambda$ attains its smallest value $\alpha(c)$ at $\lambda_o\in(0,1)$. Hence letting $\lambda=\lambda_o$ in (35), we attain from (17) that

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t)-g_t(j)Q_t(j|X_{1_t},X_{2_t})]\le\alpha(c).$$

$$\mu-a.s.\qquad \omega\in D(c). \tag{36}$$

By (7), (6), (15), (29) and (36), noticing $g_t(j)=-\log Q_t(j|X_{1_t},X_{2_t})$, we can deduce

$$\limsup_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t},X_{2_t})]$$

$$= \limsup_{n\to\infty}\{-\frac{1}{|T^{(n)}|}\log\mu(X^{T^{(n)}}) - \frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o\}} H_t^Q(X_t|X_{1_t})\}$$

$$\leq \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\{-\log Q_t(X_t|X_{1_t},X_{2_t}) - H_t^Q(X_t|X_{1_t},X_{2_t})\}$$

$$+\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}[\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log Q_t(X_t|X_{1_t},X_{2_t}) - \log\mu(X^{T^{(n)}})] = \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}$$

$$\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{j\in S}\{-\delta_j(X_t)\log Q_t(j|X_{1_t},X_{2_t}) + Q_t(j|X_{1_t},X_{2_t})\log Q_t(j|X_{1_t},X_{2_t})\}$$

$$+\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\log\frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}$$

$$= \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{j=s_1}^{s_M}[g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t},X_{2_t})]$$

$$+\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\log\frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}$$

$$le\sum_{j=s_1}^{s_M}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t},X_{2_t})]$$

$$\leq \alpha(c)M \qquad \mu-a.s. \qquad \omega\in D(c), \tag{37}$$

thus in the case $c > 0$, (19) follows from (37).
In the case $-1 < \lambda < 0$, by (31) we have

$$\liminf_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t},X_{2_t})]$$

$$\geq \frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log^2 Q_t(j|X_{1_t},X_{2_t})\cdot Q_t(j|X_{1_t},X_{2_t})^{1+\lambda} + \frac{c}{\lambda}.$$

$$\mu-a.s. \qquad \omega\in D(c). \tag{38}$$

By (34) and (38), we gain

$$\liminf_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}[g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t},X_{2_t})]$$

$$\geq \frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}(\frac{2}{1+\lambda})^2e^{-2} + \frac{c}{\lambda}$$

$$\geq \frac{2\lambda e^{-2}}{(1+\lambda)^2}+\frac{\mathrm{c}}{\lambda}.$$

$$\mu - a.s. \qquad \omega \in D(c). \tag{39}$$

In the case $c > 0$, the function $u(\lambda) = (2\lambda e^{-2})/(1+\lambda)^2 + c/\lambda$ attains the largest value $\beta(c)$ at $\lambda^o \in (-1, 0)$. Thereby letting $\lambda = \lambda^o$ in (39), we have

$$\liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} [g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t}, X_{2_t})] \geq \beta(c).$$

$$\mu - a.s. \qquad \omega \in D(c). \tag{40}$$

By (7), (6), (13), (16) and (40), noticing that $g_t(j) = -\log Q_t(j|X_{1_t}, X_{2_t})$, we can write

$$\begin{aligned}
&\liminf_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} H_t^Q(X_t|X_{1_t}, X_{2_t})\\
&\geq \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \{-\log Q_t(X_t|X_{1_t}, X_{2_t}) - H_t^Q(X_t|X_{1_t}, X_{2_t})\}\\
&+ \liminf_{n\to\infty} \frac{1}{|T^{(n)}|}[\sum_{t\in T^{(n)}\setminus\{o,-1\}} \log Q_t(X_t|X_{1_t}, X_{2_t}) - \log \mu(X^{T^{(n)}})]\\
&\geq \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \sum_{j=s_1}^{s_M} [g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t}, X_{2_t})]\\
&- \limsup_{n\to\infty} \frac{1}{|T^{(n)}|}[\log \mu(X^{T^{(n)}}) - \sum_{t\in T^{(n)}\setminus\{o\}} \log Q_t(X_t|X_{1_t})]\\
&\geq \sum_{j=s_1}^{s_M} \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} [g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t}, X_{2_t})] - \varphi(\mu|\mu_Q)\\
&\geq \beta(c)M - c. \qquad \mu - a.s. \qquad \omega \in D(c).
\end{aligned} \tag{41}$$

In accordance with (41), we see that (20) also holds in the case $c > 0$. When $c = 0$, take $0 < \lambda_i < 1, (i = 1, 2, \cdots)$ such that $\lambda_i \to 0 \ (i \to \infty)$, by (35) we acquire

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} [g_t(j)\delta_j(X_t) - g_t(j)Q_t(j|X_{1_t}, X_{2_t})] \leq 0.$$

$$\mu - a.s. \qquad \omega \in D(0). \tag{42}$$

Imitating the proof of (37), we have by (17) and (42)

$$\limsup_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} H_t^Q(X_t|X_{1_t}, X_{2_t})] \leq 0. \ \mu - a.s. \qquad \omega \in D(0). \tag{43}$$

Since $\alpha(0)=0$, we know that (19) also holds in the case $c=0$ from (37). By the similar means, we can obtain that (20) holds in the case $c=0$.

Corollary 1. *Under the assumption of Theorem 1, we have that in the case* $0\le c<1$,

$$\limsup_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t},X_{2_t})]\le M[\frac{2e^{-2}}{(1-\sqrt{c})^2}+1]\sqrt{c},$$

$$\mu-a.s. \qquad \omega\in D(c) \tag{44}$$

$$\liminf_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t},X_{2_t})]\ge -M[\frac{2e^{-2}}{(1-\sqrt{c})^2}+1]\sqrt{c}-c.$$

$$\mu-a.s. \qquad \omega\in D(c). \tag{45}$$

Proof. Letting $x=\sqrt{c}$ in (19) and (20), we have

$$[2e^{-2}/(1-\sqrt{c})^2+1]\sqrt{c}\ge\alpha(c), \quad -[2e^{-2}/(1-\sqrt{c})^2+1]\sqrt{c}\le\beta(c).$$

Therefore (44), (45) follow from (19) and (20), respectively.

Corollary 2. *Under the assumption of Theorem 1, we have*

$$\lim_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t},X_{2_t})]=0. \quad \mu-a.s. \qquad \omega\in D(0). \tag{46}$$

Proof. Letting $c=0$ in Corollary 1, (46) follows from (44) and (45).

Corollary 3. *Let $X=\{X_t,t\in T\}$ be the second-order nonhomogeneous Markov chains field indexed by the double rooted tree with the initial distribution (5) and the joint distribution (6), $f_n(\omega)$ be defined as (8). Then*

$$\lim_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t},X_{2_t})]=0. \qquad \mu_Q-a.s. \tag{47}$$

Proof. Let $\mu\equiv\mu_Q$ in Theorem 1, then $\varphi(\mu|\mu_Q)\equiv 0$. Thereby $D(0)=\Omega$. (47) follows from (46) correspondingly.

Remark. When the second-order nonhomogeneous Markov chain indexed by the tree degenerates into the first-order nonhomogeneous Markov chain indexed by a tree, we can see easily that $Q_t(X_t|X_{1_t},X_{2_t})=Q_t(X_t|X_{1_t})$, $H_t^Q(X_t|X_{1_t},X_{2_t})=H_t^Q(X_t|X_{1_t})$. At the moment, (47) is changed into

$$\lim_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_t^Q(X_t|X_{1_t})]=0. \qquad \mu_Q-a.s.$$

This is a main result of Yang and Ye (see [13]).

3. Shannon-McMillan theorem for homogeneous Markov chains fields on a double rooted tree

Let $X=\{X_t, t\in T\}$ be another second-order homogeneous Markov chain indexed by a double rooted tree with the initial distribution and the joint distribution on the measure μ_P as follows:

$$\mu_P(x_0,x_{-1})=p(x_0,x_{-1}) \tag{48}$$

$$\mu_P(x^{T^{(n)}})=p(x_0,x_{-1})\prod_{t\in T^{(n)}\setminus\{o,-1\}}P(x_t|x_{1_t},x_{2_t}),\quad n\geq 1, \tag{49}$$

where $P=P(z|x,y),\ x,y,z\in S$ is a strictly positive stochastic matrix on S^3, $p=(p(x,y))$ is a strictly positive distribution on S^2. Thereby the relative entropy density of $X=\{X_t, t\in T\}$ on the measure μ_P is

$$f_n(\omega)=-\frac{1}{|T^{(n)}|}[\log p(X_0,X_{-1})+\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log P(x_t|x_{1_t},x_{2_t})]. \tag{50}$$

Let a be a real number, denote $[a]^+=\max\{a,0\}$. We have the following result:

Theorem 2. *Let $X=\{X_t, t\in T\}$ be the second-order homogeneous Markov chains field with the initial distribution (48) and joint distribution (49) under the measure μ_P. $f_n(\omega)$ is defined by (50). Let $Q=Q(z|x,y),\ x,y,z\in S$ be defined by definition 1, $\alpha=\min\{Q(j|i,k),\ i,k,j\in S\}>0$. Denote*

$$H_Q(X_t|X_{1_t},X_{2_t})=-\sum_{x_t\in S}Q(x_t|X_{1_t},X_{2_t})\log Q(x_t|X_{1_t},X_{2_t}),\quad t\in T^{(n)}\setminus\{o,-1\}.$$

If

$$\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}[P(j|i,k)-Q(j|i,k)]^+\leq \alpha\cdot c, \tag{51}$$

then

$$\limsup_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_Q(X_t|X_{1_t},X_{2_t})]\leq \alpha(c)M. \qquad \mu_P-a.s. \tag{52}$$

$$\liminf_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}H_Q(X_t|X_{1_t},X_{2_t})]\geq \beta(c)M-c. \qquad \mu_P-a.s. \tag{53}$$

Proof. Let $\mu=\mu_P,\ Q_t(z|x,y)\equiv Q(z|x,y),\ x,y,z\in S,\ t\in T^{(n)}\setminus\{o,-1\}$, we obtain that $H_t^Q(X_t|X_{1_t},X_{2_t})=H_Q(X_t|X_{1_t},X_{2_t}),\ t\in T^{(n)}\setminus\{o,-1\}$, thus (50) follows from (7) and (49). By the inequalities $\log x \le x-1$ $(x>0)$, $a \le [a]^+$ and (51), noticing that $\alpha=\min\{Q(j|i,k),\ i,k,j\in S\}>0$, we can conclude

$$\begin{aligned}
\varphi(\mu_P|\mu_Q) &= \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\log\frac{\mu_P(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})}\\
&= \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\log\frac{p(X_0,X_{-1})\prod_{t\in T^{(n)}\setminus\{o,-1\}}P(X_t|X_{1_t},X_{2_t})}{q(X_0,X_{-1})\prod_{t\in T^{(n)}\setminus\{o,-1\}}Q(X_t|X_{1_t},X_{2_t})}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\log\frac{p(X_0,X_{-1})}{q(X_0,X_{-1})}+\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\log\frac{P(X_t|X_{1_t},X_{2_t})}{Q(X_t|X_{1_t},X_{2_t})}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\delta_j(X_t)\delta_i(X_{1_t})\delta_k(X_{2_t})\log\frac{P(j|i,k)}{Q(j|i,k)}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\delta_j(X_t)\delta_i(X_{1_t})\delta_k(X_{2_t})\frac{P(j|i,k)-Q(j|i,k)}{Q(j|i,k)}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\frac{P(j|i,k)-Q(j|i,k)}{Q(j|i,k)}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\frac{[P(j|i,k)-Q(j|i,k)]^+}{\alpha}\\
&\le \sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}\frac{[P(j|i,k)-Q(j|i,k)]^+}{\alpha}\\
&\le \sum_{i\in S}\sum_{k\in S}\sum_{j\in S}\limsup_{n\to\infty}\frac{|T^{(n)}|-2}{|T^{(n)}|}\frac{[P(j|i,k)-Q(j|i,k)]^+}{\alpha}\\
&= \frac{1}{\alpha}\sum_{i\in S}\sum_{k\in S}\sum_{j\in S}[P(j|i,k)-Q(j|i,k)]^+. \qquad (54)
\end{aligned}$$

By (51) and (54) we have

$$\varphi(\mu_P|\mu_Q)\le c. \qquad a.s. \qquad (55)$$

By (16) and (55) we know $D(c)=\Omega$. Hence (52), (53) follow from (19), (20), respectively.

Theorem 3. *Let $X=\{X_t,t\in T\}$ be a second-order homogeneous Markov chains field indexed by the double rooted tree with the initial distribution (1) and the transition matrix $Q=Q(z|x,y),\ x,y,z\in S$. Then*

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{o,-1\}}E_Q[H_Q(X_t|X_{1_t},X_{2_t})]=-\sum_{i\in S}\sum_{j\in S}\sum_{k\in S}q(i,j)Q(k|i,j)\log Q(k|i,j),$$

$$\mu_Q - a.s., \tag{56}$$

where E_Q represents the expectation under the measure μ_Q.

Proof. By the definition of $H_Q(X_t|X_{1_t}, X_{2_t})$ in Theorem 2 and the property of the conditional expectation, we can write

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} E_Q[H_Q(X_t|X_{1_t}, X_{2_t})] \\
= &\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} E_Q[E_Q(-\log Q(X_t|X_{1_t}, X_{2_t})|X_{1_t}, X_{2_t})] \\
= &\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} E_Q(-\log Q(X_t|X_{1_t}, X_{2_t})) \\
= &\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \sum_{x_{1_t}\in S} \sum_{x_{2_t}\in S} \sum_{x_t\in S} [-q(x_{1_t}, x_{2_t}, x_t) \log Q(x_t|x_{1_t}, x_{2_t})] \\
= &\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \sum_{i\in S} \sum_{j\in S} \sum_{k\in S} [-q(i,j,k) \log Q(k|i,j)] \\
= &\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o,-1\}} \sum_{i\in S} \sum_{j\in S} \sum_{k\in S} [-q(i,j)Q(k|i,j) \log Q(k|i,j)] \\
= &\sum_{i\in S} \sum_{j\in S} \sum_{k\in S} [-q(i,j)Q(k|i,j) \log Q(k|i,j)] \cdot \lim_{n\to\infty} \frac{|T^{(n)}| - 2}{|T^{(n)}|} \\
= &\sum_{i\in S} \sum_{j\in S} \sum_{k\in S} [-q(i,j)Q(k|i,j) \log Q(k|i,j)].
\end{aligned} \tag{57}$$

Therefore, (56) follows from (57) directly.

4. Acknowledgements

Authors would like to thank the referee for his valuable suggestions. The work is supported by the Project of Higher Schools' Natural Science Basic Research of Jiangsu Province of China (13KJB110006). Wang Kangkang is the corresponding author.

References

[1] I. Benjammini and Y. Peres, Markov chains indexed bu trees. Ann. Probab. 22(2): 219-243. 1994.

[2] T. Berger, and Z. Ye, Entropic aspects of random fields on trees. IEEE Trans. Inform. Theory. 36(3): 1006-1018. 1990.

[3] R. Pemantle, Antomorphism invariant measure on trees. Ann. Probab. 20(3):1549-1566. 1992

[4] Z.X. Ye, and T. Berger, Ergodic regularity and asymptotic equipartition property of random fields on trees. J.Combin.Inform.System.Sci, 21(1):157-184. 1996.

[5] Z.X. Ye, and T. Berger, Information Measures for Discrete Random Fields. Science Press, New York. 1998.

[6] W.G. Yang Some limit properties for Markov chains indexed by homogeneous tree . Stat. Probab. Letts. 65(2): 241-250. 2003

[7] W. Liu, and W.G. Yang, An extension of Shannon-McMillan theorem and some limit properties for nonhomogeneous Markov chains. Stochastic Process. Appl, 61(1), 129-145. 1996

[8] W.G. Yang and W. Liu, Strong law of large numbers and Shannon-McMillan theorem for Markov chains fields on trees. IEEE Trans.Inform.Theory. 48(1): 313-318. 2002

[9] W.C. Peng, W.G. Yang and B. Wang, A class of small deviation theorems for functionals of random fields on a homogeneous tree. Journal of Mathematical Analysis and Applications. 361, 293-301. 2010

[10] K.K. Wang and D.C. Zong Some Shannon-McMillan approximation theoems for Markov chain field on the generalized Bethe tree. Journal of Inequalities and Applications. Article ID 470910, 18 pages doi:10.1155/2011/470910 2011

[11] H.L. Huang and W.G. Yang Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree, Science in China, 51(2): 195-202 2008.

[12] J.L. Doob Stochastic Process. Wiley, New York. 1953.

[13] W.G. Yang and Z.X. Ye The asymptotic equipartition property for nonhomogeneous Markov chains indexed by a homogeneous tree. IEEE Trans.Inform.Theory. 53(5): 3275-3280. 2007

Wang Kangkang
School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China
E-mail: wkk.cn@126.com

Zong Decai
Department of Computer Science and Engineering, Changshu Institute of Technology, Changshu 215500, China

Received 11 April 2012
Accepted 17 October 2013